

Stabilization time for a type of evolution on binary strings

Jacob Funk ¹

Mihai Nica ²

Michael Noyes

Abstract

We consider a type of evolution on $\{0, 1\}^n$ which occurs in discrete steps whereby at each step, we replace every occurrence of the substring “01” by “10”. After at most $n - 1$ steps we will reach a string of the form $11 \cdots 1100 \cdots 00$, which we will call a “stabilized” string. If we choose each bit of the string independently to be a 1 with probability p and a 0 with probability $1 - p$, then the time of stabilization of a string in $\{0, 1\}^n$ is a random variable with values in $\{0, 1, \dots, n - 1\}$. We study the asymptotic behavior of this random variable as $n \rightarrow \infty$ and we determine its limit distribution after suitable centering by mean and scaling by variance. The case $p = \frac{1}{2}$ turns out to be different from any other values of p ; in this case the limit distribution is χ_3 , which is distribution of the length of $(Z_1, Z_2, Z_3) \in \mathbb{R}^3$ where Z_1, Z_2, Z_3 are independent $N(0, 1)$ random variables. For $p \neq \frac{1}{2}$, the limit distribution is Gaussian $N(0, p(1 - p))$. The tools used in our arguments are a natural interpretation of strings in $\{0, 1\}^n$ as Young diagrams, and a connection with the known distribution for the maximal height of a Brownian path on $[0, 1]$.

1 Introduction

For $n \in \mathbb{N}$ and $p \in (0, 1)$ let Ω_n^p denote the probability space consisting of strings in $\{0, 1\}^n$, where elements ω are assigned the product probability induced by assigning each bit independently to be a 1 with probability p or a 0 with probability $1 - p$. We consider the following kind of ‘evolution’ for strings in Ω_n^p : for a fixed $\omega \in \Omega_n^p$ we look at all the occurrences of the substring “01” in ω and we replace every one of them by “10”. By doing so, new instances of “01” may be created (for instance $0101 \mapsto 1010$ creates a “01” in the middle). We repeat this process until we reach a string of the form $11 \cdots 1100 \cdots 00$.

A concrete example: say we have $n = 8$ and we start with the string $\omega = 01101011$. Then our evolution produces the following strings before stabilizing:

$$01101011 \mapsto 10110101 \mapsto 11011010 \mapsto 11101100 \mapsto 11110100 \mapsto 11111000,$$

and stabilizes after 5 iterations because there are no more instances of “01” to be found.

This evolution has a cute interpretation, the original inspiration for the problem, as a line of confused soldiers, see [5]. There is also an interpretation as particles whose motion is restricted to one dimension. Imagine that each 1 in the string is a particle that would like to move to the left-hand side of the string and that each 0 is an empty space. At every iteration, if a particle has an empty space to its left, it will move into that space. This is exactly the replacement rule $01 \rightarrow 10$. Equally well, one can think of the 0’s as the particles which are trying to move as far right as they can and the 1’s as open spaces. The process

¹Research supported by an USRA Award from NSERC, Canada.

²Research supported by a MacCracken fellowship from New York University

will stabilize when all of the particles have moved as far as they can go. In this guise, the problem is a kind of deterministic analogue of certain "exclusion processes" (see for instance [3] Chapter 8) except in our case the initial condition is random, but the evolution is deterministic.

It is elementary to show that this process must stabilize after at most $n - 1$ iterations, and that afterwards we will obtain a string of the form $11 \cdots 1100 \cdots 00$. The number of iterations until we reach such a final configuration is a random variable on the probability space Ω_n^p . We will call this random variable the *stabilization time* and denote it by $T_n^p : \Omega_n^p \rightarrow \{0, 1, \dots, n - 1\}$. Going back to our concrete example above, we have $T_8^p(01101011) = 5$.

In this paper we will examine the limit distribution for the random variable T_n^p in the limit $n \rightarrow \infty$ and for varying values of p . One of the points of interest is the fact that the limit distribution in the case $p = \frac{1}{2}$ is qualitatively different than it is for other values of p .

Theorem 1.1. *We have the following weak limits for the distribution of the random variable T_n^p in the limit $n \rightarrow \infty$:*

In the case $p < \frac{1}{2}$:

$$\frac{T_n^p - (1 - p)n}{\sqrt{n}} \Rightarrow N(0, p(1 - p))$$

In the case $p = \frac{1}{2}$:

$$\frac{T_n^p - \frac{1}{2}n}{\sqrt{n}} \Rightarrow \chi_3$$

In the case $p > \frac{1}{2}$:

$$\frac{T_n^p - pn}{\sqrt{n}} \Rightarrow N(0, p(1 - p))$$

In the above limits, $N(0, p(1 - p))$ is a mean zero Gaussian variable with variance $p(1 - p)$, and $\chi_3 \sim \sqrt{Z_1^2 + Z_2^2 + Z_3^2}$ is the Euclidean norm of a vector of three independent standard $N(0, 1)$ Gaussian variables. This has density:

$$d\chi_3(x) = \frac{8\sqrt{2}}{\sqrt{\pi}} x^2 e^{-2x^2} dx \text{ for } x > 0$$

The proof of the theorem comes through connecting several ideas from combinatorics and probability theory. First, we notice a natural connection between strings in Ω_n^p and Young diagrams. Using Young diagrams as a tool, we can analyze the special case of strings that begin with a 0 and end with a 1 and relate the stabilization time to a simple random walk in one dimension. Using this connection to the random walk, we can find the limit distribution for the stabilization time in the special case mentioned above using the central limit theorem and Donsker's theorem. Finally, we show that the limit distribution for the special case of strings is actually the same as the limit distribution for general strings.

We have divided the proof into three lemmas which are stated below and each discussed and proved in their own sections.

Lemma 1.2. *Given a string $\omega \in \{0, 1\}^n$, there is a natural 1D random walk associated with ω which takes a step up for every 0 in ω and takes a step down for every 1 in ω . To be precise, for $0 \leq k \leq n$ let $S_k = \sum_{i=1}^k (1 - 2\omega_i)$. In the special case where $\omega_1 = 0$ and $\omega_n = 1$, we have the following explicit relationship between the random walk S_k and the stabilization time $T_n^p(\omega)$:*

$$T_n^p(\omega) = \frac{n}{2} + \max_{1 \leq k \leq n} S_k - \frac{S_n}{2} - 1$$

Remark 1.3. The proof of this lemma comes by mapping each string to a Young diagram in a natural way. The stabilization time of string turns out to be equal to the property called the *depth* of the Young diagram. In the special case $\omega_1 = 0$ and $\omega_n = 1$, the depth of the Young diagram is also found to be equal to the above expression for our random walk. The proof of this lemma is discussed in Section 2.

Lemma 1.4. *Let $\tilde{\Omega}_n^p = \{\omega \in \{0, 1\}^n : \omega_1 = 0, \omega_n = 1\}$ be the probability space where each bit, except for the first and last, is chosen independently at random to be a 1 with probability p and to be 0 with probability $1 - p$. The stabilization time of a string is a random variable on this probability space which we will denote by $\tilde{T}_n^p : \tilde{\Omega}_n^p \rightarrow \{1, 2, \dots, n - 1\}$. \tilde{T}_n^p has the following weak limits as $n \rightarrow \infty$:*

In the case $p < \frac{1}{2}$:

$$\frac{\tilde{T}_n^p - n(1 - p)}{\sqrt{n}} \Rightarrow N(0, p(1 - p))$$

In the case $p = \frac{1}{2}$:

$$\frac{\tilde{T}_n^p - \frac{1}{2}n}{\sqrt{n}} \Rightarrow \chi_3$$

In the case $p > \frac{1}{2}$:

$$\frac{\tilde{T}_n^p - np}{\sqrt{n}} \Rightarrow N(0, p(1 - p))$$

Remark 1.5. Once $\tilde{T}_n^p(\omega) = \frac{n}{2} + \max_{1 \leq k \leq n} S_k - \frac{S_n}{2} - 1$ is established in Lemma 1.2, the proof of Lemma 1.4 is an exercise using the central limit theorem and Donsker's theorem. For $p < \frac{1}{2}$ some simple analysis shows that the term $\max_{1 \leq k \leq n} S_k$ does not contribute to the limit distribution, and the central limit theorem is enough to prove the limit. The case $p > \frac{1}{2}$ follows by symmetry between the 0's and 1's in the problem. In the case of $p = \frac{1}{2}$, the random variable χ_3 arises from a connection to Brownian motion. More precisely, if B_t is a Brownian motion and $M_t = \max_{s \leq t} B_s$ is its running maximum then:

$$M_1 - \frac{1}{2}B_1 \stackrel{d}{=} \chi_3$$

and for a simple random walk S_k , we have by Donsker's theorem that:

$$\frac{\max_{1 \leq k \leq n} S_k - \frac{1}{2}S_n}{\sqrt{n}} \Rightarrow M_1 - \frac{1}{2}B_1$$

The full details of the proof of this lemma are displayed in Section 3.

Lemma 1.6. *Let $\tilde{\mu}_r^p$ be the law of the special case random variable \tilde{T}_r^p and let μ_n^p be the law of the random variable T_n^p . Let δ_0 be the unit mass at 0.*

For $p = \frac{1}{2}$ these measures are related by:

$$\mu_n^{\frac{1}{2}} = \frac{n+1}{2^n} \delta_0 + \sum_{r=0}^{n-2} \frac{n-r-1}{2^{n-r}} \tilde{\mu}_{r+2}^{\frac{1}{2}}$$

For $p \neq \frac{1}{2}$ these measures are related by:

$$\mu_n^p = \frac{p^{n+1} - (1-p)^{n+1}}{2p-1} \delta_0 + \sum_{r=0}^{n-2} \frac{(1-p)p^{n-r} - p(1-p)^{n-r}}{2p-1} \tilde{\mu}_{r+2}^p$$

Moreover, from this relation, we can show that the random variable T_n^p will have the same limit distribution as the special case random variable \tilde{T}_n^p .

Remark 1.7. The formulas stated in Lemma 1.6 are obtained by carefully counting the number of ways of adding leading 1's on the left and trailing 0's on the right of a string of special type $\tilde{\omega} \in \tilde{\Omega}_n^p$, and by exploiting the obvious fact that such additions of leading 1's and trailing 0's does not affect stabilization time. Hence, the stabilization time for a general string ω is equal to the stabilization time for the special case substring $\tilde{\omega}$ one obtains from ω by deleting any leading 1's or trailing 0's. This gives the relationship between $\tilde{\mu}$ and μ in the lemma. One can see that this is a convex combination of the probability measures $\tilde{\mu}_r$, which is heavily weighted towards higher values of r . Some elementary analysis is used to show that this weighting is such that μ and $\tilde{\mu}$ have the same limit distribution. The details of this lemma are given in Section 4.

2 Connection to Random Walk in the Special Case

In this section we aim to prove Lemma 1.2 which gives the following explicit formula for the stabilization time of $\omega \in \{0, 1\}^n$ satisfying $\omega_1 = 0$ and $\omega_n = 1$.

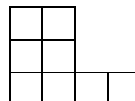
$$T_n^p(\omega) = \frac{n}{2} + \max_{1 \leq k \leq n} S_k - \frac{S_n}{2} - 1$$

This formula, once the statement is known, can be proved directly by induction. However, the proof by induction obscures the mechanics of what is really happening in the evolution. A more illuminating solution, which is how the formula was first discovered, is to map the strings into Young diagrams in a particular way. When viewed as Young diagrams in this way, the evolution process becomes particularly simple, and then the explicit formula can be seen more directly. This is the proof we will present in this section. We will begin by creating a map from $\{0, 1\}^n$ to the space of Young diagrams.

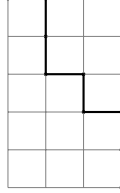
Definition 2.1. Let \mathbb{Y} denote the collection of all finite sets $Y \subseteq \mathbb{N} \times \mathbb{N}$ which have the following property.

$$(i, j) \in Y \Rightarrow (i', j') \in Y \text{ for all } i', j' \in \mathbb{N} \text{ such that } i' \leq i \text{ and } j' \leq j.$$

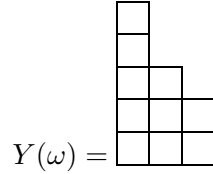
The empty set is also counted in \mathbb{Y} . It is customary to represent a set $Y \in \mathbb{Y}$ as a picture containing a “collection of boxes” (square boxes of side 1), where for every $(i, j) \in Y$ we take in our picture the box which has (i, j) as its top right corner. In this guise, the set Y is called a *Young diagram*, see e.g. [4]. The following example illustrates this representation.

$$Y = \{(1, 1), (2, 1), (3, 1), (4, 1), (1, 2), (2, 2), (1, 3), (2, 3)\} =$$


Definition 2.2. For every $\omega \in \{0,1\}^n$ with $\omega_1 = 0$ and $\omega_n = 1$, let $U = U(\omega)$ be the number of 1's in the string ω . Let B be a $U \times (n - U)$ grid on the x-y plane. Reading ω from left to right, we construct a path starting from the top-left corner of B by drawing a line horizontally to the right whenever we encounter a 0 in ω and a line vertically downward whenever we encounter a 1 in ω . Since there are U “1”'s and $n - U$ “0”'s to be found in ω we will get a path from the top-left corner in B to the bottom-right corner in B . Here is an example of the path generated by the string $\omega = (0, 1, 1, 0, 1, 0, 1, 1)$:



Define $\pi(\omega) \subset \mathbb{N} \times \mathbb{N}$ to be the unique path constructed in this manner. In our example, $\pi(0, 1, 1, 0, 1, 0, 1, 1) = \{(0, 5), (1, 5), (1, 4), (1, 3), (2, 3), (2, 2), (3, 2), (3, 1), (3, 0)\}$. The set of boxes under this path defines a Young diagram, which we denote $Y(\omega)$. In our example, this is the following diagram.



Definition 2.3. Let $Y \in \mathbb{Y}$ be a Young diagram and let $(i, j) \in Y$. We say that (i, j) is an *exposed corner* of Y if $(i + 1, j) \notin Y$ and $(i, j + 1) \notin Y$. The *corner cutting map* $K : \mathbb{Y} \rightarrow \mathbb{Y}$ is the map that removes all the exposed corners of a Young diagram as follows.

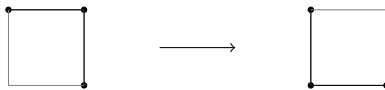
$$K(Y) := Y \setminus \{(i, j) \in Y \mid (i, j) \text{ is an exposed corner of } Y\}$$

It is easily verified by the definition that removing the exposed corners does indeed yield another Young diagram. That is to say, K is a well-defined function from $\mathbb{Y} \rightarrow \mathbb{Y}$.

Proposition 2.4. Let $\omega \in \{0,1\}^n$ be a string and let ω' be the string obtained from ω after one stage of the evolution, that is after replacing instances of “01” with “10” once. Then the Young diagram $Y(\omega')$ is obtained by cutting the corners of the Young diagram $Y(\omega)$.

$$Y(\omega') = K(Y(\omega))$$

Proof. Consider the path $\pi(\omega)$ defined in Definition 2.2. Any instances of “01” in ω will correspond to a horizontal segment followed by a vertical segment, while instances of “10” correspond to a vertical segment followed by a horizontal segment. As such, the evolution “01” \rightarrow “10 will translate into the following pictorial evolution for the path $\pi(\omega)$.



These instances of “01” are correspond exactly to the exposed corners of $Y(\omega)$ since they have no neighbors above them or to their right. We can see that this evolution is precisely removing these exposed corners of $Y(\omega)$. ■

Definition 2.5. Let $Y \in \mathbb{Y}$, $Y \neq \emptyset$ be a non-empty Young diagram. We define:

$$\text{Depth}(Y) = \max\{i + j - 1 \mid (i, j) \in Y\}$$

We also set the convention $\text{Depth}(\emptyset) = 0$.

Proposition 2.6. Let $Y \in \mathbb{Y}$, $Y \neq \emptyset$. Then $\text{Depth}(K(Y)) = \text{Depth}(Y) - 1$.

Proof. Suppose $(i, j) \in Y$ such that $i + j - 1 = \text{Depth}(Y)$. Clearly $(i + 1, j) \notin Y$ and $(i, j + 1) \notin Y$, since assuming otherwise contradicts $i + j - 1 = \text{Depth}(Y)$ is maximal. Then (i, j) is an exposed corner, so $(i, j) \notin K(Y)$. Hence $\text{Depth}(K(Y)) < \text{Depth}(Y)$.

Now, if either $(i - 1, j) \in K(Y)$ or $(i, j - 1) \in K(Y)$ then $\text{Depth}(K(Y)) \geq i + j - 2 = \text{Depth}(Y) - 1$. By combining the inequalities, we conclude $\text{Depth}(K(Y)) = \text{Depth}(Y) - 1$. Otherwise, both $(i - 1, j) \notin K(Y)$ and $(i, j - 1) \notin K(Y)$. But this only happens in the case $Y = \{(1, 1)\}$, in which case $K(Y) = \emptyset$ and the proposition holds by $\text{Depth}(\emptyset) = 0$. ■

Lemma 2.7. Let $\omega \in \{0, 1\}^n$ and let $T(\omega)$ be its stabilization time. Then we have the following relation.

$$T(\omega) = \text{Depth}(Y(\omega))$$

Proof. The proof follows by induction on $\text{Depth}(Y(\omega))$ using the previous two propositions. For the base case, if $\text{Depth}(Y(\omega)) = 0$, then $Y = \emptyset$ so ω is already stable and the stabilization time is 0 and the result holds. Now assume that for all ω with $\text{Depth}(Y(\omega)) = k - 1$, $T(\omega) = k - 1$. Given any ω with $\text{Depth}(Y(\omega)) = k$, let ω' be the evolution of ω by one step. Then, by the last two propositions we have that $\text{Depth}(Y(\omega')) = \text{Depth}(K(Y(\omega))) = \text{Depth}(Y(\omega)) - 1 = k - 1$, and so by the induction hypothesis, $T(\omega') = k - 1$. Since we have applied one step to get from ω to ω' , $T(\omega) = T(\omega') + 1 = k$. ■

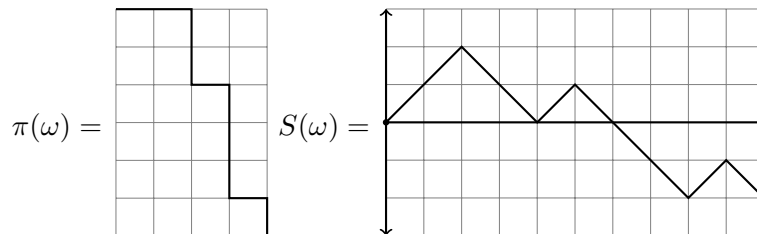
Proposition 2.8. Let $\omega \in \{0, 1\}^n$, and for $0 \leq k \leq n$ let $S_k = \sum_{i=1}^k (1 - 2\omega_i)$. Let $U = U(\omega)$ be the number of 1's in the string ω . In the case that $\omega_1 = 0$ and $\omega_n = 1$ we have the following explicit relationship between S_k and the depth of the Young diagram $Y(\omega)$:

$$\text{Depth}(Y(\omega)) = U + \max_{0 \leq k \leq n} S_k - 1$$

Proof. Let $\pi(\omega)$ be the path associated to ω in Definition 2.2. An elementary computation shows that S is related to $\pi(\omega)$ by the following formula:

$$\{(k, S_k) : 0 \leq k \leq n\} = \{(i + (U - j), i - (U - j)) : (i, j) \in \pi(\omega)\}.$$

This formula merely says that (because of the specifics of how each of S and $\pi(\omega)$ are constructed from ω) the set of lattice points $S(\omega) := \{(k, S_k) : 0 \leq k \leq n\}$ is obtained out of $\pi(\omega)$ via 45° degree rotation and dilation by $\sqrt{2}$. The verification of the formula is left as exercise to the reader. An illustration of how $\pi(\omega)$ and $S(\omega)$ look in a concrete case is shown in the next picture, drawn for $\omega = (0, 0, 1, 1, 0, 1, 1, 1, 0, 1) \in \{0, 1\}^{10}$.



Now, if $\omega_1 = 0$ and $\omega_n = 1$, we know that the boundary of the Young diagram $Y(\omega)$ is precisely the path $\pi(\omega)$. (The fact that $\omega_1 = 0$ and $\omega_n = 1$ is needed here because otherwise there are some points in $\pi(\omega)$ which are not included in $Y(\omega)$). Since the $\text{Depth}(Y(\omega))$ is achieved somewhere on its boundary, we have that:

$$\begin{aligned} \text{Depth}(Y(\omega)) &= \max \{i + j - 1 : (i, j) \in Y(\omega)\} \\ &= \max \{i + j - 1 : (i, j) \in \pi(\omega)\} \\ &= U + \max \{i - (U - j) : (i, j) \in \pi(\omega)\} - 1 \\ &= U + \max \{S_k : 0 \leq k \leq n\} - 1 \end{aligned}$$

The last equality follows from the map between S and $\pi(\omega)$ described above, and the complete equality is our desired result. ■

From here the result of Lemma 1.2 is a simple corollary obtained by using the elementary fact that $U = \frac{1}{2}(n - S_n)$.

3 Limit Distribution in the Special Case

In this section we aim to prove the aforementioned weak limits for the random variable \tilde{T}_n^p , the stabilization time for strings from the probability space $\tilde{\Omega}_n^p = \{\omega \in \{0, 1\}^n : \omega_1 = 0, \omega_n = 1\}$. We will use the result from Lemma 1.2 connecting \tilde{T}_n^p to the random walk associated with ω , $S_k = \sum_{i=1}^k (1 - 2\omega_i)$. Recall that:

$$\tilde{T}_n^p(\omega) = \frac{n}{2} + \max_{1 \leq k \leq n} S_k - \frac{S_n}{2} - 1$$

Lemma 3.1. *Let X_1, X_2, \dots, X_{n-2} be i.i.d random variables which take the value -1 with probability p and 1 with probability $1 - p$. For $0 \leq k \leq n - 2$, let $W_k = \sum_{i=1}^k X_i$ be the random walk which takes the X_i 's as its steps. Then:*

$$\tilde{T}_n^p \stackrel{d}{=} \frac{n}{2} + \max_{0 \leq k \leq n-2} W_k - \frac{W_{n-2}}{2}$$

Proof. Make the identification that $X_i \stackrel{d}{=} (1 - 2\omega_{i+1})$. Since $\omega_1 = 0$, $\omega_n = 1$, we have that $S_{k+1} \stackrel{d}{=} 1 + W_k$ for $0 \leq k \leq n - 2$ and $S_n = 1 + W_{n-2} - 1$. From this, it is clear that $\max_{1 \leq k \leq n} S_k = 1 + \max_{0 \leq k \leq n-2} W_k$, and the result is then immediate from Lemma 1.2. ■

Since W_k is the sum of many i.i.d. random variables, we are in a position to use tools like the central limit theorem and Donsker's theorem to find the limit distribution. We divide the remaining results into the case when $p \neq \frac{1}{2}$ and the case when $p = \frac{1}{2}$.

Proposition 3.2. *For $p > \frac{1}{2}$, we have that, as $n \rightarrow \infty$:*

$$\frac{\max_{0 \leq k \leq n-2} W_k}{\sqrt{n}} \xrightarrow{\mathbf{P}} 0$$

Proof. $\max_{0 \leq k < \infty} W_k$ is the maximum height achieved at any time by a weighted random walk, which takes steps upwards with probability $1 - p < \frac{1}{2}$ and steps downwards with

probability $p > \frac{1}{2}$. It is a result from elementary probability that this is distributed like a *geometric random variable* related to the parameter $q = \frac{1 - \sqrt{1 - 4p(1-p)}}{2p} < 1$:

$$\max W_k \stackrel{d}{=} \text{Geom}(1 - q) - 1$$

Here, q is the probability that if the random walk W_k reaches height x , the walk will *ever* reach the height $x + 1$. The random walk can either reach $x + 1$ immediately by taking a step upward at this time step (with probability $1 - p$) or it can step down now, and then take two steps upward eventually (these are independent, so this happens with probability $p \cdot q \cdot q$). Hence q satisfies the quadratic equation $q = (1 - p) + pq^2$, the solving of which admits the stated value for q . To see that the maximum is indeed distributed as a geometric random variable, write:

$$\begin{aligned} \mathbf{P}(\max W_k = m) &= \mathbf{P}(\exists k : W_k = 1 | \exists i : W_i = 0) \cdot \mathbf{P}(\exists k : W_k = 2 | \exists i : W_i = 1) \cdot \dots \\ &\quad \dots \cdot \mathbf{P}(\forall k, W_k \neq m + 1 | \exists i : W_i = m) \\ &= q \cdot q \cdot \dots \cdot q \cdot (1 - q) = q^m (1 - q) \end{aligned}$$

from which the relationship to the geometric random variable is clear. In particular, $\mathbf{E}(\max W_k) = \frac{1}{1-q} - 1 < \infty$ and so we have by the Markov inequality that:

$$\begin{aligned} \mathbf{P}\left(\frac{\max_{0 \leq k \leq n-2} W_k}{\sqrt{n}} > \epsilon\right) &\leq \frac{\mathbf{E}(\max_{0 \leq k \leq \infty} W_k)}{\epsilon \sqrt{n}} \\ &\rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

■

Lemma 3.3. *Let $\tilde{\Omega}_n^p = \{\omega \in \{0, 1\}^n : \omega_1 = 0, \omega_n = 1\}$ be a probability space and \tilde{T}_n^p be the stabilization time on this probability space both as defined in Lemma 1.4. For $p > \frac{1}{2}$, \tilde{T}_n^p has the following weak convergence to a Gaussian as $n \rightarrow \infty$:*

$$\frac{\tilde{T}_n^p - np}{\sqrt{n}} \Rightarrow N(0, p(1 - p))$$

Proof. We have, using the results of Lemma 3.1 and Proposition 3.2, that:

$$\begin{aligned} \frac{\tilde{T}_n^p - np}{\sqrt{n}} &= \frac{\left(\tilde{T}_n^p - \frac{n}{2}\right) - \frac{n}{2}(2p - 1)}{\sqrt{n}} \\ &\stackrel{d}{=} \frac{\left(\max_{0 \leq k \leq n-2} W_k - \frac{W_{n-2}}{2}\right) - \frac{n}{2}(2p - 1)}{\sqrt{n}} \\ &= \frac{\max_{0 \leq k \leq n-2} W_k}{\sqrt{n}} - \frac{1}{2} \frac{W_{n-2} - n(1 - 2p)}{\sqrt{n}} \\ &\Rightarrow 0 + Z_{p(1-p)} \end{aligned}$$

Since $\frac{\max_{0 \leq k \leq n-2} W_k}{\sqrt{n}} \Rightarrow 0$ by Proposition 3.2 and because W_{n-2} is the sum of the i.i.d. random variables X_i with mean $1 - 2p$ and variance $4p(1 - p)$, so by the Central Limit Theorem, we have weak convergence to a Gaussian: $\frac{W_{n-2} - n(1 - 2p)}{\sqrt{n}} \Rightarrow N(0, 4p(1 - p))$. ■

Corollary 3.4. *The same limit distribution as in Lemma 3.3 holds for $p < \frac{1}{2}$.*

Proof. As described in the introduction, we can imagine the evolution process from two points of view: we can think of each 1 in the string is a particle trying to move to the left of the string, or we could equally well imagine that each 0 is a particle trying to move to the right of the string. Each of these ways of thinking about the problem is equally valid, and gives the correspondence between p and $1 - p$.

More formally, one can identify the probability space $\tilde{\Omega}_n^p$ with the probability space $\tilde{\Omega}_n^{1-p}$ with the following bijection: interchanging the 0's and 1's and then reverse the order of the string. This bijection preserves the evolution operation of “01” \rightarrow “10” and hence shows that that \tilde{T}_n^p and \tilde{T}_n^{1-p} have the same distribution for every n . ■

We now state “Donsker’s Theorem”, which is used to find the limit distribution in the case $p = \frac{1}{2}$. – see e.g. Section 8 of [1].

Theorem 3.5. (*Donsker’s Theorem.*) *Let X_1, X_2, X_3, \dots be i.i.d. random variables that are ± 1 each with probability $\frac{1}{2}$. Let $W_n = \sum_{1 \leq j \leq n} X_j$ be the sum up to time n . (This is a simple symmetric 1D random walk). For each n , define a piecewise linear function $L^n(t)$ that is linear on each interval $[\frac{k-1}{n}, \frac{k}{n}]$ so that $L^n(\frac{k}{n}) = \frac{W_k}{\sqrt{n}}$. Then, we have weak convergence of the law of $L^n(\cdot)$ to the law of a Brownian motion $B_t, t \in [0, 1]$ in the sense of weak convergence in the space of continuous functions $C[0, 1]$ equipped with the sup-norm metric:*

$$L^n(\cdot) \Rightarrow B_{(\cdot)}$$

Lemma 3.6. *Let $B_t, t \in [0, 1]$ be a Brownian motion and let $M_t = \max_{s \leq t} B_s$ be it’s running maximum. Then, in the case that $p = \frac{1}{2}$ we have the following weak limit:*

$$\frac{\tilde{T}_n^{\frac{1}{2}} - \frac{n}{2}}{\sqrt{n}} \Rightarrow M_1 - \frac{1}{2}B_1$$

Proof. We know that $\tilde{T}_n^p \stackrel{d}{=} \frac{n}{2} + \max_{0 \leq k \leq n-2} W_k - \frac{W_{n-2}}{2}$, so it suffices to show that:

$$\frac{\max_{0 \leq k \leq n} W_k - \frac{1}{2}W_n}{\sqrt{n}} \Rightarrow M_1 - \frac{1}{2}B_1$$

This is a consequence of Donsker’s theorem. Indeed, let $h : C[0, 1] \rightarrow \mathbb{R}$ by $h(f(\cdot)) = \sup_{t \in [0, 1]} f(t) - \frac{1}{2}f(1)$. This is a continuous function on $C[0, 1]$ with the sup norm (indeed, both $\sup_{t \in [0, 1]} f(t) = \|f\|$ and evaluation at time 1 are continuous in this space). Now, since h is continuous, we know that it respects weak convergence. Applying our h to the limit in Donsker’s theorem, $L^n(\cdot) \Rightarrow B_{(\cdot)}$ gives $h(L^n(\cdot)) \Rightarrow h(B_{(\cdot)})$. By our definition of h and of L^n , this is precisely that:

$$\begin{aligned} h(L^n(\cdot)) &= \sup_{t \in [0, 1]} L^n(t) - \frac{1}{2}L^n(1) \\ &= \frac{\max_{0 \leq k \leq n} W_k - \frac{1}{2}W_n}{\sqrt{n}} \\ &\Rightarrow h(B_{(\cdot)}) \\ &= \sup_{t \in [0, 1]} B_t - \frac{1}{2}B_1 \\ &= M_1 - \frac{1}{2}B_1 \end{aligned}$$

■

Lemma 3.7. $M_1 - \frac{1}{2}B_1 \stackrel{d}{=} \chi_3$ with probability density function:

$$d\chi_3 = \frac{8\sqrt{2}}{\sqrt{\pi}} x^2 \exp(-2x^2) dx$$

Proof. We verify this by computing the density of $M_1 - \frac{1}{2}B_1$. This is just a computation using the joint density for Brownian motion and its Maximum, which is readily calculated using the reflection principle – see for example [2] pg 95. The joint density function for Brownian motion and its maximum is:

$$\rho(M_T = b, B_T = a) = \frac{2(2b - a)}{\sqrt{2\pi}T^{\frac{3}{2}}} \exp\left(-\frac{(2b - a)^2}{2T}\right)$$

for $b > a, b > 0$, and 0 otherwise. Now to get the density function for $M_1 - \frac{1}{2}B_1$, one just integrates the joint density for B_t and M_t along a line:

$$\begin{aligned} \rho(H = x) &= \int_{-2x}^{2x} \rho\left(M_1 = \frac{y}{2} + x, B_1 = y\right) dy \\ &= \sqrt{\frac{2}{\pi}} \int_{-2x}^{2x} ((y + 2x) - y) \exp\left(-\frac{((y + 2x) - y)^2}{2}\right) dy \\ &= \sqrt{\frac{2}{\pi}} \int_{-2x}^{2x} (2x) \exp\left(-\frac{(2x)^2}{2}\right) dy \\ &= \frac{8\sqrt{2}}{\sqrt{\pi}} x^2 \exp(-2x^2) \end{aligned}$$

■

Combining all the results from this section gives the result of Lemma 1.4.

4 Limit Distribution in the General Case

So far, we have results in the special case that the string ω has $\omega_1 = 0$ and $\omega_n = 1$. In this section, we will bootstrap off of these results to see that we have the same limit in the general case where there is no restriction on ω_1 or ω_n .

Lemma 4.1. Let $\tilde{\mu}_r^p$ be the law of the special case random variable \tilde{T}_r^p and let μ_n^p be the law of the random variable T_n^p . For $p \neq \frac{1}{2}$ the measures are related by:

$$\mu_n^p = \frac{p^{n+1} - (1-p)^{n+1}}{2p-1} \delta_0 + \sum_{r=0}^{n-2} \frac{(1-p)p^{n-r} - p(1-p)^{n-r}}{2p-1} \tilde{\mu}_{r+2}^p$$

Proof. We begin by splitting the space $\{0, 1\}^n$ into disjoint subsets. For stable $\omega \in \{0, 1\}^n$, all “1”s in ω must lie to the left of all “0”s in ω . As such, there are $n + 1$ stable strings in $\{0, 1\}^n$ for which the time to stabilization is zero. For each $1 \leq i \leq n$, there is precisely one such string with exactly i “1”s. These strings contribute the following value to μ_n^p .

$$\sum_{i=0}^n p^i (1-p)^{n-i} \delta_0 = \frac{p^{n+1} - (1-p)^{n+1}}{2p-1} \delta_0 \quad (4.1)$$

For non-stable $\omega \in \{0, 1\}^n$, we will introduce r , the number of elements of ω lying between the first “0” and the last “1”, so that we may write $\omega = (1, \dots, 1, 0, x_1, \dots, x_r, 1, 0, \dots, 0)$, with $x \in \{0, 1\}^r$. The time until stabilization of ω is the time until stabilization of $(0, x_1, x_2, \dots, x_r, 1)$. For given r , the distribution of these times is precisely distributed like \tilde{T}_{r+2}^p because this string is in the special case. It remains to count the number of ways we can get r to be a fixed value. Outside of (x_1, x_2, \dots, x_r) , we require one “1”, one “0”, and $n - 2 - r$ instances of either “1” or “0” in arrangements that are unique for a given number of “1”s and “0”s. These strings contribute the following value to μ_n^p for a given r :

$$\sum_{i=0}^{n-r-2} p^{i+1} (1-p)^{n-r-i-1} \tilde{\mu}_{r+2}^p = \frac{(1-p)p^{n-r} - p(1-p)^{n-r}}{2p-1} \tilde{\mu}_{r+2}^p \quad (4.2)$$

By summing over all possible values of r , we get the desired result. \blacksquare

Lemma 4.2. *When $p = \frac{1}{2}$, we have:*

$$\mu_n^{\frac{1}{2}} = \frac{n+1}{2^n} \delta_0 + \sum_{r=0}^{n-2} \frac{n-r-1}{2^{n-r}} \tilde{\mu}_{r+2}^{\frac{1}{2}}$$

Proof. The proof is the same as above. The only difference is that in the case $p = \frac{1}{2}$ the sums we need to evaluate are arithmetic sums, instead of geometric ones. One can also obtain the result by taking the limit $p \rightarrow \frac{1}{2}$ and computing the indeterminate limits. \blacksquare

Notation 4.3. We introduce the notation $C_{r,n}^p$, a set of positive coefficients defined by the results of Lemma 4.1 and 4.2, so that the following holds true for all values of p :

$$\mu_n^p = C_{0,n}^p \delta_0 + \sum_{r=2}^n C_{r,n}^p \tilde{\mu}_r^p$$

The reason for doing this is to unify the notation from the cases $p = \frac{1}{2}$ and $p \neq \frac{1}{2}$. To prove the result we are after, we only need three properties of the coefficients $C_{r,n}^p$. These are proven in the next lemma.

Lemma 4.4. *For each n , we choose a parameter θ_n , for which $\theta_n \rightarrow \infty$ as $n \rightarrow \infty$ and $\frac{\theta_n}{\sqrt{n}} \rightarrow 0$ as $n \rightarrow \infty$. We will use $\theta_n = \lfloor n^\alpha \rfloor$ where $0 < \alpha < \frac{1}{2}$ as an illustration of these properties. Given θ_n as defined above, we have the following properties of $C_{r,n}^p$:*

$$\begin{aligned} 1. \quad & C_{0,n}^p + \sum_{r=2}^n C_{r,n}^p = 1 \\ 2. \quad & C_{0,n}^p + \sum_{r=2}^{n-\theta_n} C_{r,n}^p \xrightarrow{n \rightarrow \infty} 0 \\ 3. \quad & \sum_{r=n-\theta_n+1}^n C_{r,n}^p \xrightarrow{n \rightarrow \infty} 1 \end{aligned}$$

Proof. Property 1 holds since the $C_{r,n}^p$ s arose in Lemma 4.1 and Lemma 4.2 as the division of the probability space Ω_n^p into disjoint pieces. This can be verified directly by summing over the values for $C_{r,n}^p$. Once 1 is established, 2 and 3 are equivalent. To establish 2, we use the definitions of $C_{r,n}^p$ which are found in Lemma 4.1 and Lemma 4.2. For $p < \frac{1}{2}$, we have the following equation.

$$\begin{aligned} C_{0,n}^p + \sum_{r=2}^{n-\theta_n} C_{r,n}^p &= \frac{p^{n+1} - (1-p)^{n+1}}{2p-1} + \sum_{r=2}^{n-\theta_n} \frac{(1-p)p^{n-r+2} - p(1-p)^{n-r+2}}{2p-1} \\ &\leq \frac{p^{n+1}}{2p-1} \cdot 1 + \frac{1-p}{2p-1} \sum_{r=2}^{n-\theta_n} p^{n-r+2} \cdot 1 \\ &= \frac{p^{n+1} + p^{\theta_n+2} - p^{n+1}}{2p-1} \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

The last line holds since $\theta_n \rightarrow \infty$ as $n \rightarrow \infty$. The proof is similar when $p > \frac{1}{2}$. Finally, when $p = \frac{1}{2}$, we have the following equation.

$$\begin{aligned} C_{0,n}^p + \sum_{r=2}^{n-\theta_n} C_{r,n}^p &= \frac{n+1}{2^n} + \sum_{r=2}^{n-\theta_n} \frac{n-r+1}{2^{n-r+2}} \\ &= \frac{n+1}{2^n} + \left(\frac{\theta_n+2}{2^{\theta_n+1}} - \frac{n+1}{2^n} \right) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

■

Proposition 4.5. *The random variable T_n^p has the same limit distribution as that of \tilde{T}_n^p . That is:*

$$\frac{T_n^p - n(p \vee (1-p))}{\sqrt{n}} \Rightarrow \rho$$

where ρ is the limit distribution of $\frac{\tilde{T}_n^p - n(p \vee (1-p))}{\sqrt{n}}$ as established in Lemma 1.4.

Proof. For the whole proof, we fix a value of p . We will discard the superscript p in $C_{r,n}$, μ_n and $\tilde{\mu}_r$ in the remainder of the proof to simplify the notation. We assume that $p \geq \frac{1}{2}$ so that the limit in question is $\frac{T_n - pn}{\sqrt{n}}$. In the case $p < \frac{1}{2}$ the limit in question is $\frac{T_n - (1-p)n}{\sqrt{n}}$ instead, but the argument is identical.

We use the following characterization of weak convergence for distributions ρ_n : $\rho_n \Rightarrow \rho$ if and only if $\rho_n((-\infty, a]) \rightarrow \rho((-\infty, a])$ for every $a \in \mathbb{R}$ with $\rho(\{a\}) = 0$. This is true since the semi-infinite intervals $\{(-\infty, a] : a \in \mathbb{R}\}$ are a so-called *convergence-determining* class. See [1] example 2.3 pg 18.

Let $A = (-\infty, a]$, $a \in \mathbb{R}$ be an arbitrary semi-infinite interval. It will be useful in our study here to use the notation $\sqrt{n}A + pn = \sqrt{n}(-\infty, a] + pn = (-\infty, \sqrt{n}a + pn]$. Our aim is to show that:

$$\begin{aligned} \left| \mathbf{P} \left(\frac{T_n - pn}{\sqrt{n}} \in A \right) - \rho(A) \right| &= \left| \mathbf{P}(T_n \in \sqrt{n}A + pn) - \rho(A) \right| \\ &= \left| (\mu_n)(\sqrt{n}A + pn) - \rho(A) \right| \\ &\xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

Once this is established, since sets of the form $[-\infty, a)$ are convergence determining, we will have weak convergence. We proceed by dividing into three terms using the triangle inequality and the set-algebra identity that $\sqrt{n}A + pn = (\sqrt{r}A + pr) \dot{\cup} (\sqrt{r}a + pr, \sqrt{na} + pn]$:

$$\begin{aligned}
\left| \mathbf{P} \left(\frac{T_n - pn}{\sqrt{n}} \in A \right) - \rho(A) \right| &= |(\mu_n)(\sqrt{n}A + pn) - \rho(A)| \\
&= \left| \left(C_{0,n}\delta_0 + \sum_{r=2}^n C_{r,n}\tilde{\mu}_r \right) (\sqrt{n}A + pn) - \rho(A) \right| \\
&\leq \left| \left(C_{0,n}\delta_0 + \sum_{r=2}^{n-\theta_n} C_{r,n}\tilde{\mu}_r \right) (\sqrt{n}A + pn) \right| \\
&\quad + \left| \left(\sum_{r=n-\theta_n+1}^n C_{r,n}\tilde{\mu}_r \right) (\sqrt{r}A + pr) - \rho(A) \right| \\
&\quad + \left| \left(\sum_{r=n-\theta_n+1}^n C_{r,n}\tilde{\mu}_r \right) (\sqrt{r}a + pr, \sqrt{na} + pn] \right|
\end{aligned}$$

We now show that each term individual goes to zero as $n \rightarrow \infty$. Each term is handled in a separate claim.

Claim.

$$\left| \left(C_{0,n}\delta_0 + \sum_{r=2}^{n-\theta_n} C_{r,n}\tilde{\mu}_r \right) (\sqrt{n}A + pn) \right| \xrightarrow{n \rightarrow \infty} 0$$

Proof of Claim. This is immediate from the second property in 4.4, using the following relation.

$$\left| \left(C_{0,n}\delta_0 + \sum_{r=2}^{n-\theta_n} C_{r,n}\tilde{\mu}_r \right) (\sqrt{n}A + pn) \right| \leq \left| C_{0,n} \cdot 1 + \sum_{r=2}^{n-\theta_n} C_{r,n} \cdot 1 \right| \xrightarrow{n \rightarrow \infty} 0$$

□

Claim.

$$\left| \left(\sum_{r=n-\theta_n+1}^n C_{r,n}\tilde{\mu}_r \right) (\sqrt{r}A + pr) - \rho(A) \right| \xrightarrow{n \rightarrow \infty} 0$$

Proof of Claim. This follows from the results of 4.4 and from the weak convergence for \tilde{T}_n , namely $\frac{\tilde{T}_n - pn}{\sqrt{n}} \Rightarrow \rho$. Fix $\epsilon > 0$. We use the weak convergence of \tilde{T}_n applied to the set A to find an $N \in \mathbb{N}$ so large so that $|\tilde{\mu}_r(\sqrt{r}A + pr) - \rho(A)| < \epsilon$ for all $r > N - \theta_N + 1$. Then, for all $n > N$ we have that:

$$\begin{aligned}
\left| \sum_{r=n-\theta_n+1}^n C_{r,n} \tilde{\mu}_r(\sqrt{r}A + pr) - \rho(A) \right| &\leq \left| \sum_{r=n-\theta_n+1}^n C_{r,n} (\tilde{\mu}_r(\sqrt{r}A + pr) - \rho(A)) \right| \\
&\quad + \left| \left(C_{0,n} + \sum_{r=2}^{n-\theta_n} C_{r,n} \right) \rho(A) \right| \\
&\leq \left(\sum_{r=n-\theta_n+1}^n C_{r,n} \right) \epsilon + \left(C_{0,n} + \sum_{r=2}^{n-\theta_n} C_{r,n} \right) \rho(A) \\
&\rightarrow 1 \cdot \epsilon + 0 \cdot \rho(A) = \epsilon
\end{aligned}$$

where the limits are from 4.4. Since this tends to ϵ as $n \rightarrow \infty$, it is eventually less than 2ϵ for sufficiently large n . Since this holds for all $\epsilon > 0$, we have the result of the claim. \square

Claim.

$$\left| \left(\sum_{r=n-\theta_n+1}^n C_{r,n} \tilde{\mu}_r \right) (\sqrt{r}a + pr, \sqrt{n}a + pn) \right| \xrightarrow{n \rightarrow \infty} 0$$

Proof of Claim. Fix an $\epsilon > 0$. Since the measure ρ has no atoms, choose a $\delta > 0$ so small so that $\rho(a, a + \delta] < \epsilon$. Now, since $\frac{\tilde{T}_n - pn}{\sqrt{n}} \Rightarrow \rho$, choose N so large so that for all $n > N - \theta_N + 1$ we have:

$$\left| \mathbf{P} \left(\frac{\tilde{T}_n - pn}{\sqrt{n}} \in (a, a + \delta] \right) - \rho(a, a + \delta] \right| < \epsilon$$

Now choose M so large, so that the following holds whenever $n > M$ and $n \geq r \geq n - \theta_n + 1$:

$$a \frac{\sqrt{n} - \sqrt{r}}{\sqrt{r}} + p \frac{n - r}{\sqrt{r}} < \delta$$

Such an M always exists because the inequalities $n \geq r \geq n - \theta_n + 1$ give:

$$\begin{aligned}
a \frac{\sqrt{n} - \sqrt{r}}{\sqrt{r}} + p \frac{n - r}{\sqrt{r}} &= a \left(\sqrt{\frac{n}{r}} - 1 \right) + p \frac{n - r}{\sqrt{r}} \\
&\leq |a| \left(\sqrt{\frac{1}{1 - \frac{\theta_n}{n}}} - 1 \right) + p \frac{\theta_n}{\sqrt{n - \theta_n}} \\
&\approx |a| \frac{1}{2} \left(\frac{\theta_n}{n} \right) + p \left(\frac{\theta_n}{\sqrt{n}} \right)
\end{aligned}$$

So we see that the limit works because of our choice of θ_n so that θ_n satisfies $\frac{\theta_n}{\sqrt{n}} \rightarrow 0$.

For n larger than both N and M then, we will have then for every r with $n \geq r \geq n - \theta_n + 1$ that:

$$\begin{aligned}
\tilde{\mu}_r(\sqrt{r}a + pr, \sqrt{n}a + pn) &= \mathbf{P} \left(\frac{\tilde{T}_r - pr}{\sqrt{r}} \in \left(a, a + \left(a \frac{\sqrt{n} - \sqrt{r}}{\sqrt{r}} + p \frac{n - r}{\sqrt{r}} \right) \right] \right) \\
&\leq \mathbf{P} \left(\frac{\tilde{T}_r - pr}{\sqrt{r}} \in (a, a + \delta] \right) \\
&\leq \rho(a, a + \delta] + \epsilon \\
&\leq \epsilon + \epsilon = 2\epsilon
\end{aligned}$$

Hence, for such n we have: $\left| \left(\sum_{r=n-\theta_n+1}^n c_{r,n} \tilde{\mu}_r \right) (\sqrt{r}a + pr, \sqrt{n}a + pn) \right| \leq 1 \cdot 2\epsilon$. Since this holds for every $\epsilon > 0$, we get the desired result. ■

References

- [1] P. Billingsley. *Convergence of Probability Measures*, Wiley, 1999.
- [2] I. Karatzas and S.E. Shreve *Brownian Motion and Stochastic Calculus*, Springer, 1998.
- [3] T.M. Liggett. *Interacting Particle Systems*, Springer, 2005.
- [4] R.P. Stanley. *Enumerative Combinatorics*, volume 2, Cambridge University Press, 1999.
- [5] *University of Michigan Undergraduate Mathematics Competition 13*, 1996, Problem 1.
www.math.lsa.umich.edu/undergrad/UMUMC/UMUMC13.pdf

“A large (but finite) number of soldiers are arranged in an east-west line, and all the soldiers are facing north. The commander shouts “Right face!” One second later, all the soldiers ought to be facing east, but they have not completely mastered “right” and “left”, so some are facing east and some west. Any soldier who is face-to-face with his neighbor realizes that there was a mistake and turns 180 degrees (disregarding the possibility that the mistake might have been the neighbor’s). One second later, when all these 180 degree turns have been completed, any soldier who is now face-to-face with a neighbor turns 180 degrees (even if he had just turned at the previous step). The process repeats in the same manner. Prove that it stops after finitely many steps.”

Jacob Funk
 Department of Pure Mathematics, University of Waterloo,
 Waterloo, Ontario N2L 3G1, Canada
 Email: jjfunk@uwaterloo.ca

Mihai Nica
 Courant Institute of Mathematical Sciences
 New York University
 251 Mercer Street New York, N.Y. 10012-1185 Mail Code: 0711
 Email: nicam@cims.nyu.edu

Michael Noyes
 Department of Pure Mathematics, University of Waterloo,
 Waterloo, Ontario N2L 3G1, Canada
 Email: mnoyes@uwaterloo.ca